Summarization of Complex Causal Diagnostic Hypotheses

William J. Long, PhD, MIT Lab for Computer Science, Cambridge, MA
Shapur Naimi, MD and M. G. Criscitiello, MD
New England Medical Center and Tufts Univ. School of Medicine, Boston, MA

The Heart Disease Program produces detailed causal diagnostic hypotheses for patients with cardiovascular diseases. This poster discusses our experience with summarizing these hypotheses for the physician. The basic approach is to merge the nodes of the structure indicating causal mechanism into the more important nodes. Analysis of the results shows that to generate effective summaries the identification of syndromes is very important, the definitions of the labels need to be carefully enforced, the causality of diseases should be supported by evidence and not just probability, and the sense of causal order must be carefully preserved.

1 INTRODUCTION

The use of detailed causal models for diagnosis enables the representation of many kinds of relationships allowing the diagnostic reasoner to make important distinctions not possible with simpler associations. However, more complex diagnoses may need explanation. We addressed this need for the Heart Disease Program (HDP)* in order to evaluate the program[1].

2 HEART DISEASE PROGRAM

Over the past ten years we have been developing the HDP to assist physicians in reasoning about the diagnosis and management of patients with cardiovascular disease[1]. The program uses the patient description and a causal physiologic model to generate detailed hypotheses for the differential diagnosis. A typical hypothesis is in figure 1.

The summarization procedure we developed merges nodes representing mechanism into the most closely associated important node and uses only the links needed to establish the structure of the diagnosis. Finally syndromes are recognized and combined into single nodes. The result of summarizing the hypothesis is in figure 2.

The summarization procedure did an effective job of shrinking the hypotheses and enabled effective evaluation. However, it was clear that the reviewer's objections were sometimes a result of inadequate exposition of the hypotheses in the summaries.

*The name has been changed from the Heart Failure Program to reflect the broader domain of application.

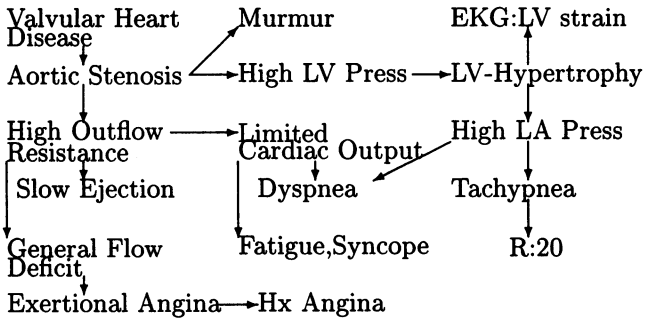

Figure 1: Fragment of a Hypothesis

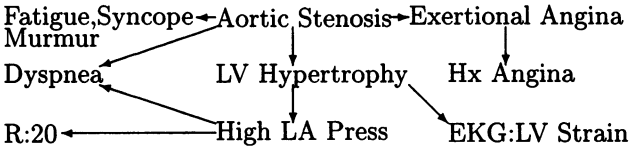

Figure 2: Summarization of Hypothesis Fragment

3 PRINCIPLES OF SUMMARIZATION

Our analysis of the summarization leads to some general principles. 1) The basic strategy of removing the mechanisms is correct. However, when only a part of the summary node causes an effect, the mechanisms are needed in a summary. 2) The sense of causality must be preserved in the summary, which may require additional knowledge. 3) The identification of syndromes is important and conveys more to the user than just the sum of the nodes. 4) Syndromes and node labels must be carefully defined to match the generally understood characteristics of acuteness, severity, or etiology. 5) A syndrome should not conceal other causes for part of the findings. 6) Unless there is evidence for a specific cause, secondary nodes or therapies should be linked to all possible causes. Nodes with a significant probability of being primary should be linked to none. 7) Mechanisms that establish a connection between nodes should remain in the summary.

References

[1] Long WJ, Naimi S, Criscitiello MG. Evaluation of a New Method for Cardiovascular Reasoning. Journal of the American Medical Informatics Association. 1994;1:127-141.

0195-4210/94/$5.00 © 1994 AMIA, Inc.